Technical Assignment 4

Ying Du (Amelia)

Faculty of Information, University of Toronto

INF2178 LEC0101

Shion Guha

March 28, 2024

Introduction

This report examines a longitudinal dataset of patients categorized into groups based on their dementia status, including those with dementia, those without dementia, and those who converted. The analysis centers on the distribution of Clinical Dementia Rating (CDR) scores across these groups and over visits.

Research Questions

After quickly went through the dataset, the research questions I'm going to focus on within this report include:

1. How do CDR scores vary between patients with and without dementia?
2. Is there a progression in CDR scores over time within patients, and how does this progression differ between groups?

Before analyzing the dataset, I cleaned it by removing all rows with missing values. Then, I performed Exploratory data analysis, mixed-effects ANOVA, a post hoc test, and power analysis for a comprehensive understanding of the research questions.

Exploratory Data Analysis (EDA)

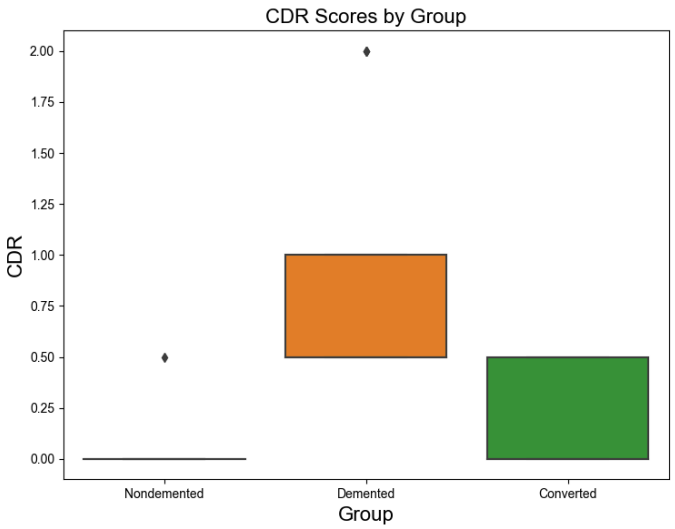

Figure 1. CDR scores across the groups

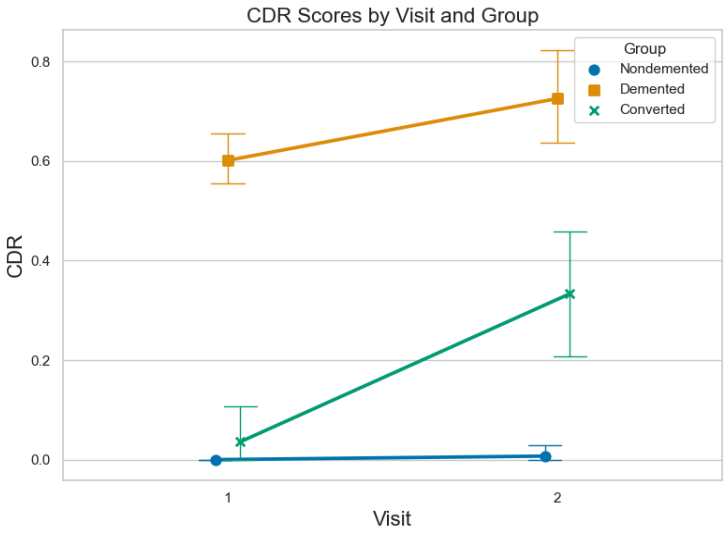

Figure 2. CDR scores by visit and groups

Figure 1 above compares CDR scores across different groups, if we focus on demented and nondemented, the interquartile ranges (IQR) clearly indicates variations in distributions. In general, patients who are categorized as demented have higher CDR scores, which indicates more severe cognitive impairment. This

visualization effectively highlights the significance of dementia status with respect to CDR scores.

To ensure statistical accuracy, I have excluded the converted group from the analysis in this report. Including the converted group in between-group analyses might introduce additional variability and potentially confound the results, as the progression of the conversion may vary widely among participants. A more comprehensive dataset would be needed to account for this variability. By focusing solely on stable groups, the analysis based on the current dataset is more likely to achieve greater and more reliable statistical results.

Looking at Figure 2 above, we can observe that patients with dementia have a higher CDR compared to those without dementia. This aligns with the information we obtained from the boxplot. Additionally, there is an increase in CDR scores for patients with dementia from the first visit to the second visit, indicating a worsening of their cognitive impairment during that period. On the other hand, for patients without dementia, there is also an increasing trend between the first and second visits, but the slope is minimal, suggesting that their CDR did not increase significantly. As a result, their cognition did not decline as much as those with dementia.

Mixed-effects ANOVA Assumptions Testing

Normality

- Shapiro-Wilk Test on CDR Scores: p-value = 1.56e-21 < 0.05

According to the Shapiro-Wilk test for normality, the CDR Scores are not normally distributed (p-value < 0.05), which does not meet the assumption of normality.

Homogeneity of Variance

- Levene's Test across 2 Groups (nondemented and demented):
  p-value = 2.01e-09 < 0.05

The Levene test for homogeneity of variances indicates that the variances of CDR scores between the nondemented and demented groups are significantly different (p-value < 0.05). This result suggests that the assumption of homogeneity of variances has been violated.

Thus, we need to keep in mind that implementing ANOVA in this dataset may affect the accuracy of the analysis result.

Mixed-effects ANOVA Result

Source	F	p-unc
Group	237.85	<0.001
Visit	19.00	<0.001
Interaction	10.48	<0.001

Table 1. ANOVA summary table

As seen in Table1, we have confirmed once again that the demented status of patients has a significant impact on their CDR scores (p-value < 0.001). We also confirmed that the CDR scores change over time for both groups of patients across two visits (p-value < 0.001). The significant interaction suggests that the pattern of change over time differs between the groups, implying that the trajectory of cognitive decline is not uniform across the different dementia statuses.

Post-hoc Test Result

Contrast	Visit	A	B	T	p-unc
Visit	-	1	2	-4.10	<0.001
Group	-	Demented	Nondemented	19.94	<0.001
Visit * Group	1	Demented	Nondemented	23.21	<0.001
Visit * Group	2	Demented	Nondemented	15.13	<0.001

Table 2. Post-hoc test results

The increase in CDR scores from Visit 1 to Visit 2 ($t = -4.10$, $p < 0.001$) across all participants signals an overall trend of cognitive deterioration over time. When comparing demented and nondemented groups, demented individuals exhibit greater cognitive decline, as indicated by significantly higher CDR scores ($t = 19.94$, $p < 0.001$). This pattern persists and indicates that patients who are demented consistently have a higher CDR ($t = 23.21$ and 15.13, $p < 0.001$) compared to nondemented participants over time.

Power Analysis Plot for T-tests

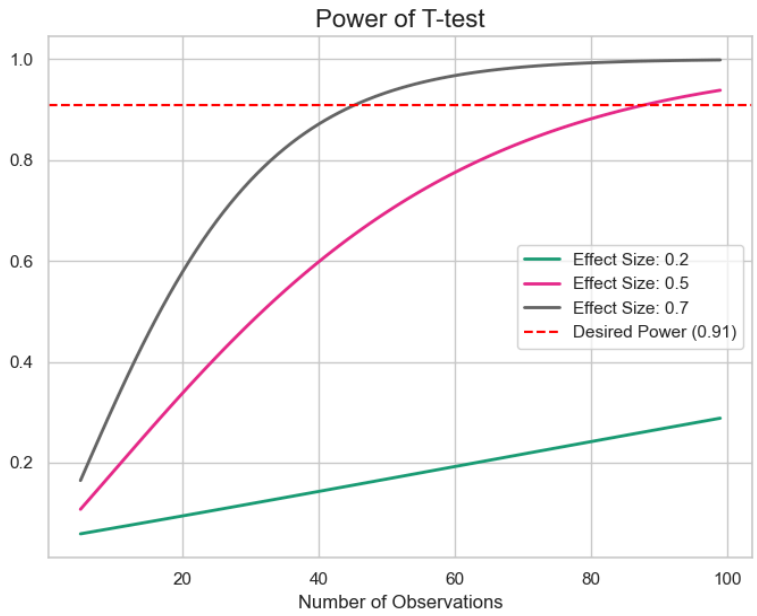

Figure 3. Power of t-Test

- Required Sample Size for Effect Size 0.7: 46

To ensure a 91% confidence level in detecting true differences in MRI scans related to dementia, my analysis suggests that we need to include MRI results from a minimum of 46 patients diagnosed with dementia and 46 patients without the condition. Figure 3 clearly represents this critical number with a horizontal dashed line, which indicates our desired power level, intersecting with the curve at the top that represents an effect size of 0.7.

Conclusion

This analysis reveals a significantly larger impact on Clinical Dementia Rating (CDR) scores for patients with dementia compared to those without, highlighting a notable impact of the condition on cognitive function. Furthermore, we observed that CDR scores progress differently over time within these groups, with a larger cognitive decline among individuals with dementia compared to those without. However, since the assumptions of mixed-effects ANOVA are not fully met, the violation could affect the validity of the results, and these findings should be interpreted with caution.

Data Reference

Harvey, P. D., & Mohs, R. C. (2001). *Clinical dementia rating*. Clinical Dementia Rating - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/neuroscience/clinical-dementia-rating